Estatística Inferencial

Prof. Wagner Hugo Bonat

Departamento de Estatística
Universidade Federal do Paraná

Intervalo de confiança

Intervalo de confiança

- **Intervalo de confiança** - Um intervalo de verossimilhança para θ é um intervalo da forma $\theta : L(\theta|\mathbf{y}) \geq r L(\hat{\theta}|\mathbf{y})$ ou equivalentemente, $\theta : D(\theta) \leq c^*$, com $D(\theta) = -2[l(\theta) - l(\hat{\theta})]$ e $c^* = -2 \log(r)$.
- r precisa ser especificado entre 0 e 1, para intervalos não vazios, logo $c^* > 0$.
- A escolha de r ou equivalentemente c^* é arbitrária.
- Já vimos que c^* pode ser atribuído olhando para a distribuição amostral.

Intervalo de confiança

- **Verossimilhança relativa**: $\frac{L(\theta)}{L(\hat{\theta})} \geq r$.
- **Função deviance**: $-2[l(\theta) - l(\hat{\theta})] \leq -2 \log(r)$.
- Problema: Encontrar $c^* = -2 \log(r)$.
- Encontrar as raízes da verossimilhança relativa ou da *deviance* em geral envolve o uso de métodos numéricos.

Função *deviance* aproximada

- Para facilitar os cálculos e obtenção de propriedades assintóticas.
- Expansão em séries de Taylor para a $l(\theta)$ em torno de $\hat{\theta}$.

$$D(\theta) = -2[l(\theta) - l(\hat{\theta})] = 2\left\{l(\hat{\theta}) - [l(\hat{\theta}) + (\theta - \hat{\theta})l'(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 l''(\hat{\theta})]\right\}.$$

- Aproximação em série de Taylor, toma a seguinte forma quadrática e define a região

$$D(\theta) = -(\theta - \hat{\theta})^2 l''(\hat{\theta}) \leq c^*.$$

que por sua vez, define intervalos de confiança da forma,

$$\hat{\theta} \pm \sqrt{\frac{c^*}{-l''(\hat{\theta})}}.$$

Exemplo: Distribuição exponencial

- Considere o caso $Y_i \sim Exp(\theta)$ iid.
- Função de verossimilhança

$$L(\theta) = \prod_{i=1}^{n} \theta \exp\{-\theta y_i\} = \theta^n \exp\{-\theta \sum_{i=1}^{n} y_i\} = \theta^n \exp\{-n\overline{y}\theta\}.$$

- Função de log-verossimilhança

$$l(\theta) = n \log \theta - \theta \sum_{i=1}^{n} y_i = n \log \theta - \theta n \overline{y}.$$

- Primeira derivada em relação a θ (função escore)

$$U(\theta) = n\theta^{-1} - n\overline{y}.$$

Exemplo: Distribuição exponencial

- Estimativa de máxima verossimilhança

$$\hat{\theta} = \frac{1}{\bar{y}}.$$

- Informação observado e esperada (negativo da segunda derivada)

$$I_O(\theta) = I_E(\theta) = n\theta^{-2} \quad \text{e} \quad I_O(\hat{\theta}) = I_E(\hat{\theta}) = n\hat{\theta}^{-2}.$$

Exemplo: Distribuição exponencial

- Intervalo baseado na *deviance*

$$
\begin{aligned}
\mathrm{D}(\theta) &= 2[\mathrm{l}(\hat{\theta}) - \mathrm{l}(\theta)] \\
&= 2[n \log \hat{\theta} - \hat{\theta} n \overline{y} - (n \log \theta - \theta n \overline{y})] \\
&= 2n[\log \left(\hat{\theta}/\theta \right) + \overline{y}(\theta - \hat{\theta})] \leq c^{*}.
\end{aligned}
$$

- Resultado: A função *deviance* segue uma distribuição qui-quadrado com 1 gl.
- Fixando $\alpha = 0{,}05$ toma-se o valor do quantil $c^{*} = 3{,}84$.
- Requer a solução de uma equação não-linear.

Exemplo: Distribuição exponencial

- Note que $\mathrm{V}(\hat{\theta}) = \mathrm{I_O}^{-1}(\hat{\theta}) = \hat{\theta}^2/n$, logo o intervalo de confiança é dado por

$$\hat{\theta}_L = \hat{\theta} - z_{\frac{\alpha}{2}}\hat{\theta}/\sqrt{n} \quad \text{e} \quad \hat{\theta}_U = \hat{\theta} + z_{\frac{\alpha}{2}}\hat{\theta}/\sqrt{n}.$$

- *Deviance* aproximada:

$$\mathrm{D}(\theta) \approx n\left(\frac{\theta - \hat{\theta}}{\hat{\theta}}\right)^2.$$

- Resolvendo a equação, tem-se

$$\left(\hat{\theta}_L \approx \hat{\theta}(1 - \sqrt{c^*/n}) \;,\; \hat{\theta}_U \approx \hat{\theta}(1 + \sqrt{c^*/n})\right).$$

Exemplo: Distribuição exponencial

- Intervalo baseado na verossimilhança relativa: $LR(\theta) = \frac{L(\theta)}{L(\hat{\theta})} \geq r$.
- Escolhendo valor de r comparável com r

$$
\begin{aligned}
\frac{L(\theta)}{L(\hat{\theta})} &\geq r \\
\log\left[\frac{L(\theta)}{L(\hat{\theta})}\right] &\geq \log r \\
-2[l(\theta) - l(\hat{\theta})] &\geq 2 \cdot \log r \\
-3{,}84 &\geq 2 \cdot \log r \\
r &\approx 0{,}146
\end{aligned}
$$

- Implementação computacional: `Exponencial.R`.

Testes de hipóteses

Conceitos iniciais

- **Hipótese estatística** - Chamamos de hipótese estatística qualquer afirmação acerca da distribuição de probabilidade de uma ou mais variáveis aleatórias.
- **Teste de hipótese** - Chamamos de teste de uma hipótese estatística a função de decisão $\chi \rightarrow \{a_0, a_1\}$, em que a_0 corresponde à ação de considerar a hipótese H_0, como verdadeira e a_1 corresponde à ação de considerar a hipótese H_1 como verdadeira.
- A função de decisão d divide o espaço amostral χ em dois conjuntos,

$$A_0 = \{(y_1, \ldots, y_n \in \chi; d(y_1, \ldots, y_n) = a_0\}$$

e

$$A_1 = \{(y_1, \ldots, y_n \in \chi; d(y_1, \ldots, y_n) = a_1\}$$

onde $A_0 \cup A_1 = \chi$ e $A_0 \cap A_1 = \emptyset$.

Tipos de erros e poder em testes de hipóteses

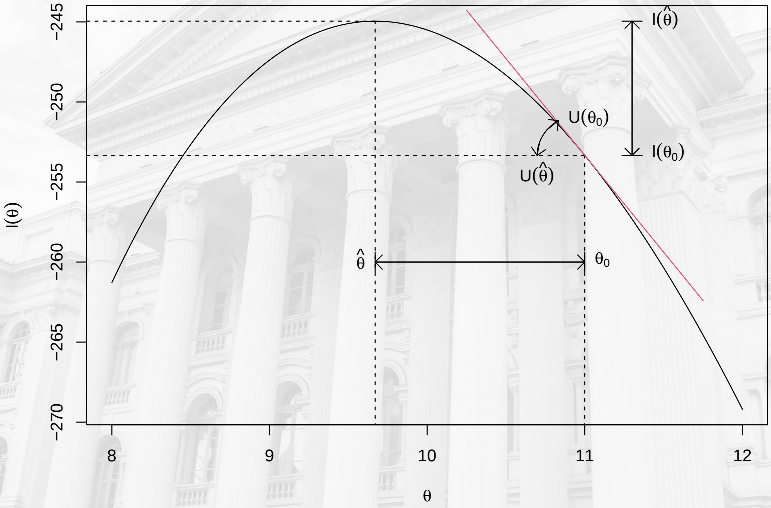
- O erro tipo I ocorre quando rejeitamos H_0 e esta é verdadeira.
- O erro Tipo II ocorre quando não rejeitamos H_0 e esta é falsa.
- Resumindo,

$$\alpha = P(Y \in A_1|\theta_0) \quad \text{e} \quad \beta = P(Y \in A_0|\theta_1).$$

- O poder do teste com região crítica A_1 para testar $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$ é dado por

$$\pi(\theta_1) = P(Y \in A_1|\theta_1).$$

Olhando para a log-verossimilhança



Teste da razão de verossimilhança

A estatística do teste da razão de verossimilhança para testar $\mathrm{H}_0 : \theta \in \Theta_0$ versus $\mathrm{H}_1 : \theta \in \Theta_0^c$ é

$$\lambda(\mathbf{y}) = \frac{\sup_{\Theta_0} \mathrm{L}(\theta|\mathbf{y})}{\sup_{\Theta} \mathrm{L}(\theta|\mathbf{y})},$$

onde sup denota o supremo de $\mathrm{L}(\theta|\mathbf{y})$ restrito ao conjunto Θ. O teste da razão de verossimilhança (TRV) é qualquer teste que tenha uma região de rejeição da forma $\mathbf{y} : \lambda(\mathbf{y}) \leq r$ onde r é qualquer número que satisfaça $0 \leq r \leq 1$.

Para testar $\mathrm{H}_0 : \theta = \theta_0$ versus $\mathrm{H}_1 : \theta \neq \theta_0$, suponha $Y_1, \ldots, Y_n$ sejam iid $f(\mathbf{y}|\theta)$, $\hat{\theta}$ seja o EMV de θ, e $f(\mathbf{y}|\theta)$ satisfaça as condições de regularidade. Desse modo, de acordo com H_0, usando resultados assintóticos à medida que $n \to \infty$

$$-2 \log \lambda(\underline{y}) \to \chi^2_1.$$

Teste Wald

- Hipótese bilateral $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.
- A estatística de Wald é dada por

$$Z_n = (\hat{\theta} - \theta_0)/\sqrt{V(\hat{\theta})}$$

  note que $Z_n \sim N(0, 1)$ quando $n \to \infty$.
- A região crítica é construída sob a hipótese nula baseada na normal padrão.

Teste escore

- A estatística de teste escore é

$$Z_S = U(\theta_0)/\sqrt{I_E(\theta_0)}.$$

Se H_0 for verdadeira, Z_S tem distribuição normal com média 0 e variância 1.

Exemplo: Distribuição Bernoulli

- Considere o caso em que $Y_i \sim B(\theta)$ iid.
- Objetivo: Testar a hipótese nula $H_0 : \theta = \theta_0$ contra a hipótese alternativa $H_1 : \theta \neq \theta_0$.
- Função de verossimilhança

$$\mathrm{L}(\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{(n-\sum_{i=1}^{n} y_i)}.$$

- Log-verossimilhança

$$\mathrm{l}(\theta) = \sum_{i=1}^{n} y_i \ln(\theta) + (n - \sum_{i=1}^{n} y_i) \ln(1-\theta).$$

Exemplo: Distribuição Bernoulli

- Função escore

$$U(\theta) = \frac{\sum_{i=1}^{n} y_i}{\theta} - \frac{n - \sum_{i=1}^{n} y_i}{(1-\theta)} = \frac{\sum_{i=1}^{n} y_i - n\theta}{\theta(1-\theta)},$$

- Estimativa de máxima verossimilhança

$$\hat{\theta} = \frac{\sum_{i=1}^{n} y_i}{n} = \overline{y}.$$

- Informação observada

$$I_O(\theta) = + \frac{\sum_{i=1}^{n} y_i}{\theta^2} + \frac{(n - \sum_{i=1}^{n} y_i)}{(1-\theta)^2}.$$

- Informação esperada

$$I_E(\theta) = E(I_O(\theta)) = \frac{n}{\theta(1-\theta)}.$$

Estatísticas de teste

- Razão de verossimilhança

$$-2\lambda(\mathbf{y}) = -2\left\{\sum_{i=1}^{n} y_i \ln\left(\frac{\theta_0}{\hat{\theta}}\right) + (n - \sum_{i=1}^{n} y_i)\ln\left(\frac{1-\theta_0}{1-\hat{\theta}}\right)\right\}.$$

- Teste Wald

$$T_W = \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \sim N(0,1).$$

- Teste escore

$$T_E = \frac{\sum_{i=1}^{n} y_i - n\theta}{\theta(1-\theta)}\sqrt{\frac{\theta(1-\theta)}{n}} = \frac{\sum_{i=1}^{n} y_i - n\theta}{\sqrt{n\theta(1-\theta)}} \sim N(0,1).$$

- Implementação computacional: `Bernoulli.R`.

Ilustração dos testes de hipóteses: Distribuição Bernoulli